**Chapter 8 Summary**

In this chapter we discussed the simplest possible qualitative response regression model in which the dependent variable is binary, taking the value of 1 if an attribute is present and the value of 0 if that attribute is absent.

Although binary dependent variable models can be estimated by OLS, in which case they are known as linear probability models (LPM), OLS is not the preferred method of estimation for such models because of two limitations, namely, that the estimated probabilities from LPM do not necessarily lie in the bounds of 0 and 1 and also because LPM assumes that the probability of a response moves linearly with the level of the explanatory variable, which is counterintuitive. One would expect the rate of increase in probability to taper off after some point.

Binary response regression models can be estimated by the logit or probit models.

The logit model uses the logistic probability distribution to estimate the parameters of the model. Although seemingly nonlinear, the log of the odds ratio, called the logit, makes the logit model linear in the parameters.

If we have grouped data, we can estimate the logit model by OLS. But if we have micro-level data, we have to use the method of maximum likelihood. In the former case we will have to correct for heteroscedasticity in the error term.

Unlike the LPM, the marginal effect of a regressor in the logit model depends not only on the coefficient of that regressor but also on the values of all regressors in the model.

An alternative to logit is the probit model. The underlying probability distribution of probit is the normal distribution. The parameters of the probit model are usually estimated by the method of maximum likelihood.

Like the logit model, the marginal effect of a regressor in the probit model involves all the regressors in the model.

The logit and probit coefficients cannot be compared directly. But if you multiply the probit coefficients by 1.81, they are then comparable with the logit coefficients. This conversion is necessary because the underlying variances of the logistic and normal distribution are different.

In practice, the logit and probit models give similar results. The choice between them depends on the availability of software and the ease of interpretation.